走馬楼呉簡吏民簿ＴＥＩテキスト化試案　　　　　　2024年9月21日　鷲尾

内容

[**一、呉簡吏民簿ＴＥＩ化の目的** 1](#_Toc185101665)

[１，簿の集成案を伝え、研究者の使用に供する。 1](#_Toc185101666)

[２，要素・属性による検索で、語彙のみならず種別等による検索・抽出が可能。 2](#_Toc185101667)

[３，情報を集計し量的研究に資す 2](#_Toc185101668)

[**二　呉簡吏民簿のＴＥＩ化について** 2](#_Toc185101669)

[**三　【四年小武陵郷簿】の書式と構成** 3](#_Toc185101670)

[**四　どのようにＴＥＩ化するか** 4](#_Toc185101671)

[**五　全体構成を踏まえて各簡を表示する** 5](#_Toc185101672)

[**六　物質としての情報、視角情報** 5](#_Toc185101673)

[１，形状（形態分類表示）、サイズ、材質、殘断情況 5](#_Toc185101674)

[２，空間配置 5](#_Toc185101675)

[**七　文字の情報** 6](#_Toc185101676)

[１，別筆　赤字の「中」字あり（簿籍/末尾(集計)/戸のみ）。＜add＞(追記・加筆を示す要素）で囲み表示、筆色はバーミリオンで表示したが、朱筆でもよいか。 6](#_Toc185101677)

[２，文字についての修正情報 6](#_Toc185101678)

[**八　内容注釈** 6](#_Toc185101679)

[**１，各簡の機能分類（様式分類）** 6](#_Toc185101680)

[**２, 様式特有の記載** 6](#_Toc185101681)

[**①戸人簡** 6](#_Toc185101682)

[**②戸成員** 7](#_Toc185101683)

[**③戸の集計** 8](#_Toc185101684)

[④里の集計・郷の総計・郷の表題 9](#_Toc185101685)

[**3, 世帯（戸）の把握** 9](#_Toc185101686)

[■結 10](#_Toc185101687)

[参考文献 10](#_Toc185101688)

# **一、呉簡吏民簿ＴＥＩ化の目的**

## １，簿の集成案を伝え、研究者の使用に供する。

　以前、ある簿の構成を詳述するために、構成簡の番号を列挙したことがある。この方法では、見る側は図版本を持っていなければ理解できない。ならば、テキストデータで簡の記載内容を列挙したほうが、わかりやすい。ＴＥＩなら、それに文字の校正情報（同じ人物が他の簿に出現する。照合するとより正確な情報が得られる）や項目の説明を織り込むことができる。また、戸の復原案を示し得る。

## ２，要素・属性による検索で、語彙のみならず種別等による検索・抽出が可能。

　ＴＥＩテキストデータは要素・属性による検索が可能である。例えば、呉簡をはじめとして公文書中の親族呼称にはバリエーションが存在し、同一の存在にたいして多数の呼称が見られることがある。属性として親族/娘のように設定すると、娘を指す多様な語を抽出可能である。簿籍に常用される集計の語には「凡」「右」などがあるが、簿籍集計の用語であることがわかるように属性を記述するとこれらの語を抽出することができる。呉簡のみならず、時代と地域の異なる簡群にわたってこれを行うことができれば、時代的変遷や地域差について考える材料となり得る。

## ３，情報を集計し量的研究に資す

　以前より、家族の様態の説明を、戸の親族構造などを通じて行ってきた。呉簡には戸の記録が大量に存在する。その各戸の完全なる復原が望ましいが難しい。戸のほかの構成員を特定できない人物も多く、そのような記述は単独で史料とするしかない。このため完全な復原結果に部分的な復原や、量的な検討を組み合わせた検討がなされてきた。量的な検討によって、戸の人数と資産は相関する傾向が明かになる(于振波2008)。また昨年発表の拙稿では、女性の初婚年齢やいつまで結婚するのか、最終的な生活の場などを復原と量的検討の組み合わせにより行った(鷲尾祐子2023a)。蘇俊林2023『身份与秩序』も田家莂記載ミスの量的な検討を行う。復原＋量的な検討によってある程度の成果が得られた。

　呉簡の最大の長所は、同じ帳簿を構成する簡が大量に残っていることにある。それは量的な検討に好適である。というよりも、集計してはじめて得られることが多いと考えられる。ＴＥＩは簿の各記載内容・項目にタグをつけることができ、属性によってそれを区別できる。要素や属性別に記載内容を抽出し集計することが容易である。ＴＥＩは量的な検討に向いている。

　もちろん、単にデータを集計するだけなら、エクセルなどにデータを入力するのが効率的である。しかし数字だけを抽出すると、もと史料の文脈や曖昧さなどが抜け落ちる（小風尚樹2022）。テキストと解釈とを両立させるＴＥＩの有効性はここに存する。また、ＴＥＩは成果の交換のため作られたので、他の研究者が量的な検討に有効利用してもらうデータになり得る。

# **二　呉簡吏民簿のＴＥＩ化について**

　呉簡には、吏民簿や戸品出錢簡（安部聡一郎2011）など簿の構成簡が特定されているものがある。これらの簿は1000枚を超える多数の簡から成立していることもある。同じ簿を構成する簡で一つのデータベースを作成したほうが、利便性が高く作成も容易である。

　同一簿の簡は同じ示意図に存在し簡冊の状態を維持しているものが多いが、ほかの簿がそこに混在していることもある。最終的にどれが同一簿なのかわかるようにマークアップするのであれば、最初から同一簿の簡を抽出してデータ化するほうが、使う側の利便性が高いのではないか。

また、量的な検討のため簿の記録をデータとして集計するにしても、同一簿の簡をいったん同じファイルにまとめることは、大量のデータの集計に向いている。異なる簿が混在していると、異なる簿の情報を排除する手間がかかる。

　また、作成も容易である。もし単に異なる簿が混在する形で簡番号順にデータ化しても、同一簿の簡はそれとわかるようにマークしていくことは可能である。しかし、種類の異なる語彙のものが混在するので、注釋などの種類が増える。構造も複雑になる。同じ書式・語彙のものが繰り返されているデータならば、コピーして内容を変えればデータ化が早い。

　文献の簡のばあい、たとえば張家山247号出土簡では 脈書と算数書の簡がごく一部混在して出土しているが、両者の簡は別の文献のものとして分けて扱う。簿籍構成簡の把握が可能な呉簡でも、おなじようにしてみてはどうだろうか。最終的には複数の簿を連結したデータを作成することを念頭に置きつつ、個別の簿のデータ作成から着手したい。

# **三　【四年小武陵郷簿】の書式と構成**

○基本的構成　いま試験的に【四年小武陵郷簿】をとりあげる。この簿は構成が単純であり、書式のゆれも無く統一されており、最も作成しやすいと考えた。各簡のサイズもみなほぼ同じである。（鷲尾祐子2023b）

〔〕はとりあえずの様式分類

（仮）様式分類は、最初の２レベルは里耶の表に従いました。それ以下は、里耶の表と合致しにくいのでひとまず独自設定しています。のちに変更する際に、機械的に一括して置換することが可能です。

１，【表題】簡冊全体の冒頭に、表題が置かれる(「小武陵郷□嘉禾四年吏民人｜名妻子年紀簿」壹10153/1・3)。〔簿籍/表題/郷〕

２【各戸の簡】続いて里ごとに各戸の簡が置かれる。

平陽里戸人公乘烝平年卅二筭一｜腫兩足　　　｜（壹10480/2・225）〔簿籍/本文/戸/戸人〕

　　｜平妻大女取年廿八筭一　　｜　　　　　　（壹10481/2・226）〔簿籍/本文/戸/戸成員〕

　　｜平子男各年七歳　　　　　｜　　　　　　（壹10488/2・234）〔簿籍/本文/戸/戸成員〕

　　｜平母大女妾年七十　　　　｜　　　　　　（壹10479/2・224）〔簿籍/本文/戸/戸成員〕

凡口三事二　｜筭二事　｜訾　五　十　（壹10380/2・129）〔簿籍/末尾(集計)/戸〕

（口は口数　その後ろの事は口錢を負担する者の人数、筭は筭銭負担対象者、事は筭銭を実際に負担する者の数。）

○ほか、疾病や傷・障碍の記載、吏卒であることを示す記載が付加されているものがある。

　東陽里戸人公乘烝敦年卌一｜筭一刑左手　｜（壹10269/1・26）〔簿籍/本文/戸/戸人)〕

高遷里戸人公乘苗覇年十七｜筭一給郡吏　｜　　　（壹10048）〔簿籍/本文/戸/戸人〕

東陽里戸人公乘謝贊年廿一筭一｜給縣卒　｜　　（壹10308）〔簿籍/本文/戸/戸人〕

○示意図の範囲内では、冒頭より吉陽里・高遷里・東陽里・平陽里の順に並ぶ。順番は不明だが、ほか宜陽里と宜都里が見える。

３，【里の集計】某里に属す各戸の簡の末尾には、里全体を集計する簡が置かれる。戸口数の集計のみ見える。

●右吉陽里領吏｜民卅六戸口食一百七十三人（壹10397/2･146）　〔簿籍/末尾(集計)/里〕

●右高遷里領吏｜民卅八戸口食一百八十人（壹10229）　〔簿籍/末尾(集計)/里〕

●右平陽里領吏｜民卅六戸口食□百□□人（壹10248/2･4）　〔簿籍/末尾(集計)/里〕

●右宜都里領｜吏民□□□口一百七人(壹4848)）。〔簿籍/末尾(集計)/里〕

４，【郷の集計】

○簡冊全体の末尾に郷全体の集計が置かれる。郷の集計として、下記の二つが見える。

〼右小武陵郷領四年吏民一百九十四［戸］、口九百五十一人、［收更口筭錢合]□□一千三百卅四錢（壹4985　凌文超2011の修正に従う）

〔簿籍/末尾(集計)/郷/総計〕

其二百五十二人筭人收錢一｜百廿合三萬二百卌(壹4980)〔簿籍/末尾(集計)/郷/内訳/筭〕

総計では、戸数総計、口数総計、更錢筭銭口錢の総計を記載。

内訳では、各錢種ごとの負担人数と合計錢数を記載

ほか更・大口・小口錢負担者人数集計があるはずだが、残存していない。

# **四　どのようにＴＥＩ化するか**

○「テキストデータは、元の媒体上でのテキストと完全に同じであるわけではない。元資料への忠実さや付与される解釈の深さは「どのような人のどのようなニーズを主要な対象とするか」「どれくらいの手間暇をかけられるか」に依拠する」（永崎研宣2022ａ）

「どのように読み・読まれるべきであり、何に着目すべきであり、どのようにして有用な情報としてそれらを取り出せるべきか」永崎2022ｂ

何に対してタグを付すか、情報の選別が必要である。、またどのように用いたいか、を考えた要素名の選択と属性の設定が望ましい。

この簿の何についてどのように符号化するか。

　機械に読ませたいこと、機械的に処理をしたいことに、タグを付す。

①物体としての情報。材質・形態分類、サイズ、残断の状況

②視覚的情報　書式による文字の配置。罫線や段。筆蹟の相違。文字の色。書体。

③様式分類

④釈文の修正と解釈の履歴。

⑤公文書中において特定の意味・機能を有する語。

量的検討のために

⑥数値の意味。

親族関係・世帯構成員の分析のために

⑥親族呼称・年齢・性別・同一戸の構成員の表示

# **五　全体構成を踏まえて各簡を表示する**

ＴＥＩガイドライン四章（訳：関西大学アジアオープンリサーチセンター東アジアＤＨポータル2022）

「番号あり区分は div1, div2, …のような名前で、この番号は階層における当該の区分の深さを表す; 最上位は ‘div1’ で、これに含まれるどんな下位区分も ‘div2’ であり、これの更なる下位区分は ‘div3’ となる (以下続く」

　郷全体に関する記述をdiv１（郷表題と郷総計

　里全体の記述をdiv２（郷集計の内訳もこの水準とした

　各戸の記述をdiv３

　段落、行などのタグもあるが、ほかの吏民簿や呉簡全体では一行一段落のみではない簡もあるので、汎用性の高いdiv（division 區文）によって一簡を表示した。

（いったんこのようにしたが、ほかの吏民簿や呉簡・他の簡牘と調整）

▲のちに、他の呉簡の簿と同一ファイルにすることを考えると、同じ簿の範囲はdiv1でくくり、その下の表題などはdiv2から、としたほうがいいかもしれない。

# **六　物質としての情報、視角情報**

## １，形状（形態分類表示）、サイズ、材質、殘断情況

　管見の諸例ではＴＥＩヘッダーにこれらの情報を記載する（書籍など、一つの物体からなるテキストを１ファイルで作成することが多い）。簡牘は簡によって異なる場合があるため、おのおのについて記載する必要がある。

　この簿のばあい材質はすべて竹なので、ヘッダーでも表示した。形態分類もすべて０１甲。

　この情報を本文中に入れると、長くなるかもしれない。最後のｂａｃｋ領域に、にリストを作成し、各簡についての形態分類・材質・サイズ・残欠情報を、簡固有情報としてまとめて記載するか。

（具体的にどの要素を用いるかわからなかった）

▲呉簡のばあい、実見調査結果のサイズと写真版サイズは異なる。写真版サイズを表示するか。

## ２，空間配置

　簿籍は罫線で区画し項目を分けて記載することがある。本簿では罫線は引かれていないが、編綴のひもがかかる部分（二箇所）によって記載する空間が分かれている（全体の表題と里簿籍/末尾(集計)/郷/総計は除外。戸人も、必ずしも拘束されない）。上段・中段・下段に分かれる。どこに書かれているか、というのはそこに何が書いているのかを知る手がかりになるため、各段を区別して表示したかった。今回は、仮に各段をセルとして表示した。

　ＴＥＩに準拠した空間配置の方法として、ほかに「layout」タグがあるようである。こちらのほうが簡易で適切であれば、こちらで記載する。

▲戸集計の冒頭「凡」は、上段の中途（だいたい上下の半分の位置）から書き始める。どうやって表示する？この簿ではすべて類似の位置から開始だが、ほかの簿ではバリエーションが存在することもある。

▲里集計は上段冒頭から三字程度下げて書き始める。こちらはどのように表示するか。

# **七　文字の情報**

## １，別筆　赤字の「中」字あり（簿籍/末尾(集計)/戸のみ）。＜add＞(追記・加筆を示す要素）で囲み表示、筆色はバーミリオンで表示したが、朱筆でもよいか。

<add rend="vermilion">中</add>

　ただし、これは加筆を意味するタグなので、別筆を意味するようにしたければ、別の要素を用いた方がよい。

▲朱筆で上から線を引いている箇所あり（10054　下段の五十のあたりから下まで太線あり　10422　下段の訾から五にかけて線状　）、これはどのように表示するか

10243・10488・4506（簡中有朱筆塗痕　10490・10497は中段から下段にかけて？）・10211（簡下有朱筆塗痕。は、写真からは不明。

## ２，文字についての修正情報

　文字の異説については、　冒頭の例に挙げた「阿佛東くだり」の例では、底本のテキストを基本的な表示とし、他の本の異文を併記するが、この形式のばあいビューワー等で底本の字が表示されるのではないか。底本の字かどうかにかかわらず、より妥当であると考えられる字を優先的に表示したい。

　このため、底本の原釋を修正したい場合には、<sic>で原釋を表示し、修正情報を <corr>で表示した。修正を提示しているのは凌文超2011なら、correspで文献リストに関連づけし、凌文超2011と記す。

○ほか

▲文字部分的にしか残っていない（つくりかへんはわかる、など部分的にわかる）場合はどうやって表示するか。

▲作字しなければ無い文字がごく少数ある。

▲原釋が推測で充当している文字は、それをどう表示するか

（これらについては、ほかの簡群のやりかたと統一したほうがいい

# **八　内容注釈**

## **１，各簡の機能分類（様式分類）**

　簡全体を囲むdivタグの中に、その區文の様式としてtype="簿籍/末尾(集計)/戸"のように記載した。

## **２, 様式特有の記載**

### **①戸人簡**

吉陽里戸人公乘張惕年廿八筭一給縣吏(壹10182/図1・32)

の記載

○里名を、<orgName type="里">吉陽里</orgName>

○戸人を <roleName type="戸人">戸人</roleName>

○爵位を　<roleName type="爵位">公乘</roleName>

○人名を　<persName corresp="#張惕10182"><surname> 張 </surname><forename>惕</forename></persName>

　名前全体を包括するpersName のタグに、correspで参照名を記入。back領域の人名リストに対応させた。

○官職名を

グラフィカル ユーザー インターフェイス, テキスト, アプリケーション

自動的に生成された説明<roleName type="官職">縣吏</roleName> （官職についてもっと細かく設定したほうがよいのではないか。）

▲凌文超2011が名前についていた？を削除して人名の釈文を確定した。履歴として記載するか。

▲年齢を＜age value="28"＞のように表示しようとしたが、ＴＥＩの規則として不可能なようだった。代替として<num type="age" value="">で表示した。

▲筭一は人頭税の対象として一名とカウントされることを示す。どのように表示するか。

▲前掲

東陽里戸人公乘烝敦年卌一｜筭一刑左手　｜（壹10269/1・26）〔簿籍/本文/戸/戸人)〕

の刑左手（左手に刀傷）のような疾病・障碍・傷など身体の状況を示す語を、

<term type="身体状況">刑左手</term>

のように表示した。意味を入れた方がよい？

▲妻・成人女性には「大女」が付される。成人女性を意味する、としてタグを付したい。何のタグで表示するか。<roleName>?

▲給は某人を吏・卒・特定の徭役に充当する際に用いる。公文書に常用される表現なので、ぜひ表示したい。ただし、解釈が分かれているので、「給」について検討しておく必要がある。

### **②戸成員**

毛(?)子男□年十苦狂病(壹10154/図1・4)

○毛は□の親。人名であることを示す。

<persName><forename>毛(?</forename></persName>

○親族呼称「子男」は、親族呼称であることを示し、日本語訳を付す。

テキスト

自動的に生成された説明<roleName type="親族/息子">子男</roleName>

### **③戸の集計**

凡口七事五　筭四事[二]　訾　一　千（壹10151/図1・1）

各集計項目について、タグをつけて集計の種類を説明し、それぞれ算用数字で数を記載した。

○口七　<num type="戸の口数" value="7">口七</num>

○事五　<num type="戸中の人頭税負担者総数" value="5">事五</num>

○筭四　<num type="戸中の筭銭負担者総数" value="4">筭四</num>

○事[二]　<num type="戸中の実際に筭銭を負担する者の総数" value="2">事[二] </num>

グラフィカル ユーザー インターフェイス が含まれている画像

自動的に生成された説明○訾　一　千　<num type="資産" value="1000"> 訾　一　千</num>

▲冒頭の「【凡】」は、秦代においてすでに簿籍の集計部分に用いられているものであり、公の簿籍にみえる特徴的な語である。これが集計に用いられることを示すよう属性を付した。属性記載はどのようにするか

<term type="簿籍/集計">凡</term>

▲最後の「事」のあとに、免除を意味する「復」が見えることがある。数値的には０として表示し得るが、「復」が負担免除を意味することを示すタグをどのように付すか。

### ④里の集計・郷の総計・郷の表題

　○三者は、段別の項目表示ではないので、表形式の表示は不要と考える（郷の集計のうち内訳簡は中段から開始しているので、表形式にしたい）。headで番号を表示し、ｐ（段落）タグで本文を表示。

　○郷はorgName districtとしてタグを付す。里名は前掲。

　○戸数・口数はさきの戸の集計と同様の方法で説明を付し、算用数字で表示する。

　○郷総計の「收更口筭錢合]□□一千三百卅四錢」は、郷の人頭税総計として説明する。

<num type="郷の人頭税総計" >收更口筭錢合]□□一千三百卅四錢</num>

▲里集計冒頭の「右」は、「凡」に同じく簿籍特有の語。合計など、それより前の記述を総括する記載に見える（【漢簡語彙　中国古代木簡辞典】縦書きの文書類で、前条を承けていう語。,『漢書』諸侯王表「右高祖十一人。呉隨父、凡十二人。」,●右卒十人　（E.P.T5:172））

ここでは、簿籍集計に用いられる語として属性を付した。

# **3, 世帯（戸）の把握**

　戸の成員複数が確定できるばあいは、戸の構成員を表示したい。

　実は、某簡と某簡は同じ戸である、という方法で表示したかったが、方法がわからなかった。そこで、back領域にリストパーソンとして個人のリストを作成し、リストリレイションで某人と某人は同じ世帯である、と表記した。グラフィカル ユーザー インターフェイス, テキスト, アプリケーション, メール

自動的に生成された説明　この方法の場合、同戸の集計簡を記載することができない。　もっと良い方法があれば、そちらを取りたい。

# ■結

　吏民簿ＴＥＩテキスト化を試みた。

○主な不明点は

▲形態分類やサイズなど物質的情報をどのようにして表示するのか

（エクセルファイル　テキストファイル、などで機械的に

簡番号をｉｄにして検索できるようにする　情報にアクセスできる。

▲原釋が推測で記載している文字はどう表示するか。[　]表記のもの。

▲公文書の特徴的な表現の記載。「右」「凡」などの集計の語彙や、「給」「大女」「筭」などの語、さらに身体的な問題点の記載の表示はどうするか。●符号の表示は？。

このほかに

▲個人の性別をどこに表示するか。名簿上では親族呼称によって性別がわかるため、男女性別は（大女以外は）特には記載しない。いまリストパーソンで性別を表示しているが、名簿本文で表示できるほうがいい。

以下は将来できたらよいと考える。

▲位置　掲剝位置示意図が存在するばあい、簡が示意図上のどの簡に対応しているのか示せると便利。

▲吏民簿中の人物は、他の簿に同一人物が見えることがある。これについても表示したほうが役に立つ。

　以上の諸案は仮のものである。より簡潔かつ適切な形式があればそちらに変更する。

また、全体の方針や、呉簡の他の簿の簡をどのように記載するかによって変更することを想定している。複数の簿にわたる集計や、書式の比較検討のために、他の史料のＴＥＩテキストと共通する記載にしたほうがいい。【四年小武陵郷簿】は書式が少なく記載内容が限られているので、あとから置換することを予定してこれを選んだ。

　いま仮に吏民簿のひとつのみについて考えた。簿一つといっても700本近くの簡からなっており、これのみでもデータを提供する価値がある。全体の仕様によって変更を余儀なくされる可能性もあるが、まず一つについて作成することには意義があると思う。

# 参考文献

永崎研宣2022ａ「人文学のためのテキストデータの構築とは」人文情報学研究所監修、石田・大向・小風・永崎・宮川・渡邉編『人文学のためのテキストデータ構築入門　ＴＥＩガイドラインに準拠した取り組みにむけて』文学通信（以下、略称『人テキ』）

永崎2022ｂ「ＴＥＩガイドラインで自分の資料を作り始めるには」『人テキ』

永崎2022ｃ「利活用演習：ＴＥＩ準拠テキストの活用方法」『人テキ』

小風尚樹2022「ＴＥＩとＲＤＦを用いた財務史料の構造化：古代日本・近世スペイン・近代イギリスを事例に」『人テキ』

関西大学アジアオープンリサーチセンター東アジアＤＨポータル　ＴＥＩガイドライン日本語訳

安部聡一郎2011「走馬楼呉簡所見“戸品出錢”簡の基礎的研究」藤田勝久・松原弘宣編『東アジア出土資料と情報伝達』汲古書院)。)

鷲尾祐子2023a「「三世紀長沙における女性のライフサイクル」『女性史学』33号

鷲尾祐子2023b「走馬楼呉簡吏民簿基本情報修正稿」東京外国語大学アジアアフリカ研究所研究課題「中国古代簡牘の横断領域的研究」ホームページ　2023年10月08日入稿

凌文超2011「走馬楼呉簡采集簡“戸籍簿”復原整理与研究——兼論呉簡“户籍簿”的類型与功能」『呉簡研究』第三輯→凌文超2015『走馬楼呉簡采集簿書整理與研究』所収。廣西師範大学出版社

蘇俊林2023『身份与秩序』広西師範大學出版社

王子今2004「三国孫呉郷村家族中的“寡嫂”和“孤兄子”——以走馬楼竹簡爲中心的考察」『古史性別研究叢稿』社会科学文献出版社

楊際平2007「秦漢戸籍管理制度研究」『中華文史論叢』2007ｰ1

于振波2008「従走馬楼呉簡看其時長沙民戸的貧富差別」『史学月刊』2008-６